

Cornell University School of Hotel Administration The Scholarly Commons

Articles and Chapters

School of Hotel Administration Collection

10-23-2009

No-Arbitrage Conditions for Storable Commodities and the Models of Futures Term Structures

Peng Liu

Cornell University School of Hotel Administration, pl333@cornell.edu

Ke Tang

Renmin University of China

Follow this and additional works at: <http://scholarship.sha.cornell.edu/articles>



Part of the [Finance and Financial Management Commons](#)

Recommended Citation

Liu, P., & Tang, K. (2009). *No-arbitrage conditions for storable commodities and the models of futures term structures* [Electronic version]. Retrieved [insert date], from Cornell University, School of Hotel Administration site: <http://scholarship.sha.cornell.edu/articles/1020>

This Article or Chapter is brought to you for free and open access by the School of Hotel Administration Collection at The Scholarly Commons. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of The Scholarly Commons. For more information, please contact hlmdigital@cornell.edu.

No-Arbitrage Conditions for Storable Commodities and the Models of Futures Term Structures

Abstract

One distinguishable feature of storable commodities is that they relate to two markets: cash market and storage market. This paper proves that, if no arbitrage exists in the storage-cash dual markets, the commodity convenience yield has to be non-negative. However, classical reduced-form models for futures term structures could allow serious arbitrages due to the high volatility of the convenience yield. To avoid negative convenience yield, this paper proposes a semi-affine arbitrage-free model, which prices futures analytically and fits futures term structures reasonably well. Importantly, our model prices commodity-related contingent claims (such as calendar spread options) quite differently with classical models.

Keywords

No-arbitrage condition, exponential affine model, convenience yield, Kalman filter

Disciplines

Finance and Financial Management

Comments

Required Publisher Statement

© [Elsevier](#). Final version published as: Liu, P., & Tang, K. (2010). No-arbitrage conditions for storable commodities and the models of futures term structures. *Journal of Banking & Finance*, 34(7), 1675-1687.

doi:10.1016/j.jbankfin.2010.03.013

Reprinted with permission. All rights reserved.

No-Arbitrage Conditions for Storable Commodities and the Models of Futures

Term Structures

Peng (Peter) Liu

Cornell University

Ke Tang*

Renmin University of China

October 23, 2009

Forthcoming in Journal of Banking & Finance

P. Liu, School of Hotel Administration, Cornell University, 465 Statler Hall, Ithaca, NY 14850, USA, pl333@cornell.edu.

K. Tang, Hanqing Advanced Institute of Economics and Finance and School of Finance, Renmin University of China, Beijing 100872, PR China, ketang@ruc.edu.cn, Tel.: +86 10 62519408; fax: +86 10 62511331.*corresponding author

Abstract

One distinguishable feature of storable commodities is that they relate to two markets: cash market and storage market. This paper proves that, if no arbitrage exists in the storage-cash dual markets, the commodity convenience yield has to be non-negative. However, classical reduced-form models for futures term structures could allow serious arbitrages due to the high volatility of the convenience yield. To avoid negative convenience yield, this paper proposes a semi-affine arbitrage-free model, which prices futures analytically and fits futures term structures reasonably well. Importantly, our model prices commodity-related contingent claims (such as calendar spread options) quite differently with classical models.

Keywords: No-arbitrage condition, exponential affine model, convenience yield,

Kalman filter

JEL Classification: G12, G13

No-Arbitrage Conditions for Storable Commodities and the Models of Futures Term Structures

1. Introduction

The stochastic behavior of commodity prices plays an important role in modeling financial contingent claims and evaluating investments to produce a commodity. However, unlike many other assets that are traded in a single market, storable commodity prices relate to two interconnected markets: the cash (financial) market and the storage (inventory) market.¹

The spot price and convenience yield serve as bridges linking the cash and storage markets. Spot prices can be considered in the same way as futures with the shortest maturity; in other words, it is the short end of the futures term structure. In the meanwhile it is also the price at which the physical commodity is traded on the storage market. To explain both the contango and backwardation futures term structures, the theory of storage (Brennan, 1958; Kaldor, 1939; Working, 1949) creates the concept of the convenience yield, which is defined through (1), where the expected return to risk-neutral investors from purchasing the commodity at t and selling it using futures for delivery at T equals the interest forgone less the convenience yield plus the storage cost:

$$F(t, T) = S_t + E_t^{\mathbb{Q}} \left[\int_t^T (S_u(r_u - \delta_u) + w) du \right], \quad (1)$$

where $F(t, T)$ denotes the futures prices observed at time t with maturity T ; S_t denotes the spot price at time t ; δ_u and r_u are, respectively, the instantaneous convenience yield and interest rate at time u , and $E_t^{\mathbb{Q}}[\cdot]$ denotes expectation up to information at t for risk-neutral investors.² w is the storage cost, which is assumed to be a constant following Brennan (1958), Fama and French

¹ For details about cash and storage markets, please refer to Pindyck (2001) and Geman and Ohana (2008).

² Eq. (1) echoes the expression for futures in Miltersen and Schwartz (1998), refer to Proposition 1 and Appendix B.

(1988) and many others. The convenience yield is a latent variable, which cannot be observed directly. Varying specifications of the convenience yield process results in quite different commodity futures behavior and contingent claim prices.

Under the framework of theory of storage and no-arbitrage pricing theory, in Section 2 of this paper we formally demonstrate that, if no arbitrage exists in the cash-storage dual markets, the convenience yield must be non-negative. The non-negativity of the convenience yield constraint is due to two features of commodities: the cash-storage dual markets feature and the no short-selling feature of physical commodities. Intuitively, if the convenience yield is negative, the futures prices will be “too high” relative to the spot price; thus an inventory trader can take so-called a “carry arbitrage” by short-selling a commodity future and then storing the physical commodity at the present time and delivering it at the futures maturity. We also show that, due to the non-negativity of the convenience yield, the slope of the futures term structure cannot exceed a certain threshold. This is consistent with many empirical studies (such as Litzenberger and Rabinowitz, 1995) where only mild contango is observed. Furthermore, in a structural equilibrium model, Routledge et al. (2000) show that as a consequence of a non-negativity constraint on inventory, the convenience yield of storable commodity represents an timing option, which must be non-negative. This is also consistent with the argument in our paper.

Most studies on commodity-related contingent claim pricing do not model the convenience yield directly. Instead they model the net convenience yield (*NCY*), which is defined as the convenience yield minus the percentage storage cost³, i.e.

$$c_t := \delta_t - \frac{w}{S_t}, \quad (2)$$

³ Since the commodity price S_t is a stochastic variable, contrary to the constant storage cost w , percentage convenience yield is a stochastic variable. Eq. (1) changes to

$$F(t, T) = S_t + E_t^{\mathbb{Q}} \left[\int_t^T (S_u(r^f + p_u - \delta_u) + w) du \right] = S_t + E_t^{\mathbb{Q}} \left[\int_t^T S_u(r^f - c_u) du \right].$$

as in Schwartz (1997), Hilliard and Reis (1998), Geman and Nguyen (2005) and Casassus and Collin-Dufresne (2005). For brevity's sake, we thus name models of this kind *NCY models*. However, storage cost, which is the sum of inventory fees charged by warehouse facilities, is usually not priced based on the spot price of stored commodities. Therefore the storage cost should in reality stay in nominal form but not in percentage terms. More importantly, if *NCY* is modeled as a Vasicek (or Ornstein–Uhlenbeck) process, it has a significant probability below negative percentage storage cost, in which case the convenience yield is negative and hence an arbitrage exists. This is mainly due to the large volatility of the convenience yield. Section 3.3 lists some empirical results about the probability of negative convenience yield.

In contrast to existing *NCY* models, we propose a three-factor (log-spot price, convenience yield and interest rates) model, where the convenience yield and storage cost are treated separately. Since the storage cost is usually fixed for a certain period and highly predictable, separately specifying the convenience yield and storage cost process is feasible. Moreover, it is easy to constrain the convenience yield from going negative if the convenience yield is treated independently. Moreover, Section 3.3 shows that separation of the convenience yield and storage cost results in a less convenience yield volatility. We emphasize the advantage of the Cox–Ingersoll–Ross model, another popular process derived from interest-rate modeling, in specifying the non-negative convenience yield process.

The usage of the nominal storage cost instead of percentage storage cost may pose a challenge in solving futures prices since it is impossible to specify an exponential affine framework, as shown in Duffie et al. (2000). However, in this paper, we propose a “semi-affine” latent factor model, which nests the traditional *NCY* type of models.⁴ We can still derive the analytical futures pricing formula. Furthermore, an extended Kalman filter is utilized to estimate

⁴ When $w = 0$, the “semi-affine” model changes to a typical affine model.

the latent factors. Using copper and oil data we show that our model fits the futures term structures reasonably well.

As shown in Lien and Yang (2008), Dempster et al. (2008), Alizadeh et al. (2008), the process of convenience yield is substantial for pricing and hedging of commodity futures and their related derivatives. As an application, we price the calendar spread option of crude oil futures under our framework. The pricing results show that our model generates lower spread option prices for out-of-the-money puts than the Schwartz (1997) model, mainly because of the non-negativity of the convenience yield. In the meanwhile, although this paper analyzes commodity futures and convenience yields, it does relate to financial futures (e.g. stock index futures). In that context, the convenience yield corresponds to the percentage dividend flow, which must be non-negative as well.⁵ Thus, models in this paper can be used to fit term structure of stock futures.

The rest of the paper is organized as follows. Section 2 formally proves that the convenience yield is non-negative in the framework of no-arbitrage pricing theory. Section 3 proposes a new three-factor model that constrains the convenience yield from going negative. Moreover, it shows that, empirically, the convenience yield has a high probability of being negative in the traditional NCY models, which indicates that the violation of no arbitrage is serious in those models. Section 4 discusses the model calibration results using oil and copper data. Section 5 proposes an application of our model for spread option pricing. Section 6 concludes.

2. No Arbitrage in Cash-Storage Dual Markets

In this section, we study no-arbitrage conditions in the cash-storage dual markets. We first derive the no-arbitrage conditions in a system consisting only of the cash market. Then, we

⁵ In this case, the storage cost is zero.

study the no-arbitrage constraints by adding the storage market to the system. Lastly, we study the implications for asset-pricing in commodity markets.

In order to analyze the no-arbitrage conditions in cash-storage dual markets, we assume bonds $B(t, T_b)$, commodity futures $F(t, T_f)$ and a money market account M exist in the cash market; and the physical commodity is the only asset in the storage market. Also, both bond and commodity futures $B(t, T_b)$ and $F(t, T_f)$ have all maturities available for trade ($T_b > t, T_f > t$) and the money market account is determined by $M := \exp(\int_0^t r_s ds)$, where r_s is the instantaneous risk-free rate at time s . Moreover, there is n -dimensional Brownian motion governing the joint movements of spot, futures and bond prices, and the covariance matrix of all bonds and futures at any time is with rank n . Also, there is a unique market-price-of-risk process for the n -dimensional Brownian motion. We also adopt the typical assumption that any trading portfolio can be rebalanced continuously with no transition cost for the rebalancing.

2.1 Equivalent martingale measure in the cash market

We take the money market account as numeraire. Typically, an asset-pricing theorem shows that a cash-market equivalent martingale measure (or risk-neutral measure) \mathbb{Q} for futures and deflated⁶ bond prices, when futures and bonds prices in the \mathbb{Q} measure follow

⁶ Deflated, here we mean security prices deflated by the numeraire.

$$E_t^{\mathbb{Q}}[dF_t] = 0 \quad (3)$$

$$E_t^{\mathbb{Q}}[dB_t] = r_t B_t dt$$

From our assumption above, it is also easy to show that the cashmarket equivalent martingale measure (EMM) is unique (refer to Duffie, 2006). The existence of EMM is sufficient to insure no arbitrage in the cash market.

2.2 Spot price process

The link between the cash and storage markets is stated in Eq. (1) under the \mathbb{Q} measure. It addresses the price relationships between the physical commodity (traded at S_t) and its financial securities (e.g., futures).

Proposition 1. *Under the \mathbb{Q} measure, if the commodity spot prices are continuously traded, the following expressions are equivalent*

$$F(t, T) = S_t + E_t^{\mathbb{Q}} \left[\int_t^T (S_u(r_u - \delta_u) + w) du \right], \quad (4)$$

$$E_t^{\mathbb{Q}}[dS_t] = (S_t(r_t - \delta_t) + w)dt, \quad (5)$$

$$S_t = E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T (\delta_u - \frac{w}{S_u} - r_u) du \right) S_T \right]. \quad (6)$$

Proof. See Appendix B.

Note that (6) together with $F(t, T) = E_t^{\mathbb{Q}}[S_T]$ form the base of futures pricing in Miltersen and Schwartz (1998) in an arbitrage-free asset-pricing framework.

Consistent with Deaton and Laroque (1992, 1996) and Routledge et al. (2000), in the storage market we assume the commodity is stored by a group of competitive inventory traders who have access to a costly storage technology with a constant rate of storage cost w . Hence, the inventory traders have to pay out storage cost over time. The storage cost w can be considered as

the *negative* dividends rate for the physical commodity asset. Therefore, in the storage market when the commodity is stored during $[0, T]$, we are more interested in finding an EMM for the ex-storage-cost process of the spot price $\hat{S}_t := S_t - wt(t \in [0, T])$.

Proposition 2. *If the convenience yield is not zero at any time, an EMM does not exist for the deflated ex-storage-cost process \hat{S}_t in the cash-storage dual markets.*

Proof. As mentioned before, since \mathbb{Q} is the unique EMM of deflated bonds and futures, when involving the spot commodity price \hat{S}_t , we need only to check whether the deflated ex-storage-cost process $\frac{\hat{S}_t}{M_t}$ is a martingale under \mathbb{Q} . Using Ito's lemma, we obtain $E_t^{\mathbb{Q}} \left[d \left(\frac{\hat{S}_t}{M_t} \right) \right] = -(S_t \delta_t) dt$, thus, only if $\delta_t = 0$, $\frac{\hat{S}_t}{M_t}$ can be a martingale, otherwise $\frac{\hat{S}_t}{M_t}$ is not a martingale, and the EMM does not exist.

Since the convenience yield is usually not zero for industrial commodities, the EMM does not exist in the cash-storage dual markets.

2.3 Non-negativity of the convenience yield

One important feature of the storage market is that *investors cannot short-sell any physical commodity*,⁷ which is consistent with the notion in Routledge et al. (2000), according to which commodities cannot be sold before they are produced. However, before we take this notion for granted, we first show through a lemma that the convenience yield will vanish if short-selling in the storage market is allowed.

Lemma 1. *If short-selling of a physical commodity is allowed in the storage market where arbitrages do not exist, the convenience yield is zero almost everywhere, i.e., $\delta_t = 0$.*

Proof. See Appendix C.

⁷ Empirically, no mechanism exists allowing investors to borrow a physical commodity through brokers.

Lemma 1 states that, if allowing short-selling of the physical commodity, the convenience yield should not exist. In other words, constraints on short-selling of physical commodities directly result in the existence of the convenience yield. To examine this lemma more clearly, we present an arbitrage scenario when $\delta_t < 0$ in case 1 and the arbitrage scenario, when $\delta_t > 0$ (and short-selling of a physical commodity is allowed) in Appendix D.

Theorem 1. *Given that short-selling is not allowed in the storage market and commodities are storable at any time, the cash-storage dual markets is arbitrage-free only if the convenience yield δ_t is always non-negative, i.e., $\delta_t \geq 0$.*

Proof. See Appendix E.

To examine this theorem more clearly, we show a carry arbitrage scenario in which $\delta_t < 0$, in the following case.

Case 1: The convenience yield. δ_t is negative at time t for an arbitrarily small time interval Δt . At time t an inventory trader I_S can buy 1 unit of a commodity by borrowing S_t amount of money from the bank and storing it immediately. Then I_S sells one unit of a commodity future with price $F(t, t + \Delta t)$, shown in (7). At time $t + \Delta t$; I_S delivers 1 physical commodity unit to the futures buyer and receives $F(t, t + \Delta t)$. Then he/she returns $S_t r_t \Delta t$ to the bank and $w \Delta t$ of storage cost to the warehouse. It is reasonable to assume there is no mark-to-market between $t, t + \Delta t$ if Δt is sufficiently small. Note that, from (1), the futures price is

$$F(t, t + \Delta t) = E_t^{\mathbb{Q}}[S_{t+\Delta t}] = S_t + (S_t(r_t - \delta_t) + w)\Delta t. \quad (7)$$

The payoff of this strategy at $t + \Delta t$ is

$$po_{t+\Delta t} = F(t, t + \Delta t) - S_t - S_t r_t \Delta t - w \Delta t = -S_t \delta_t \Delta t > 0. \quad (8)$$

Therefore, if δ_t is negative for a period lasting from t to $t + H$, and I_S can perform carry arbitrage strategy continuously from t to $t + H$, the arbitrage payoff at $t + H$ is

$$po_{t+H} = \lim_{\Delta t \rightarrow 0} \sum_j \left(-S_{t_j} \delta_{t_j} \Delta t \right) = \int_t^{t+H} (-S_s \delta_s) ds > 0.$$

In the next section, we show the implications of this notion for asset pricing in commodity markets.

2.4 Asset pricing implications in the cash-storage dual markets

The non-negativity constraint has important implications in pricing commodity-related contingent claims. We first study, under the risk-neutral \mathbb{Q} measure, the pricing of a contingent claim C (such as an option, a swap) with bonds, futures and spot prices as underlying in the cash-storage dual markets. The payoff of contingent claim C can relate to the futures, bond and commodity spot prices at an arbitrary maturity T_C . Since the spot price S_{T_C} is equal to the maturing futures prices $S_{T_C} = F(T_C, T_C)$, the payoff of C at T_C on spot commodity prices S_{T_C} is identical with the payoff of $F(t, T_C)$ at T_C . In other words, the dependency of C on the spot price S_{T_C} is the same as it is on $F(t, T_C)$. By the same token, ultimately we can make C depend only on futures and bonds but not on spot prices. Thus, any contingent claim C can be replicated only using futures and bonds in the cash market. More importantly, it is not desirable to resort to the physical storage market to replicate C because (1) C can be fully replicated the cash market where an EMM can be found, (2) the trading strategy on a physical commodity is constrained to be long-only and hence may not be able to fully track the payoff of the contingent claim. Therefore, although usually we cannot find an EMM for the ex-storage-cost spot price in the storage market, in contingent claims pricing we can ignore the storage market and price claims as if they are connected to the cash market only.

The non-negativity of the convenience yield also influences the shape of futures term

structures. Specifically, the slope of the futures term structure cannot exceed a certain threshold.

Proposition 3. *Assuming independence between commodity spot prices and interest rates, the slope of the futures term-structure with maturity T cannot exceed the product of interest-rate futures $f(t, T)$ and the commodity futures price $F(t, T)$:*

$$\frac{\partial F(t, T)}{\partial T} \leq w + F(t, T)f(t, T). \quad (9)$$

Proof. Taking derivatives of T on (1),

$$\frac{\partial F(t, T)}{\partial T} = w + E_t^{\mathbb{Q}}[S_T(r_T - \delta_T)]$$

because of the non-negativeness of the convenience yield, the maximal slope of the futures occurs when the convenience yield δ_t is zero. Hence,

$$\sup_{\delta_T} \left\{ \frac{\partial F(t, T)}{\partial T} \right\} = w + E_t^{\mathbb{Q}}[S_T r_T].$$

Furthermore, with the assumption of independence between the spot prices and interest rates, the maximal slope of the futures term structure is

$$\sup_{\delta_T} \left\{ \frac{\partial F(t, T)}{\partial T} \right\} = w + E_t^{\mathbb{Q}}[S_T]E_t^{\mathbb{Q}}[r_T] = w + F(t, T)f(t, T).$$

Note that the independence assumption with respect to interest rates and spot prices is consistent with assumptions made in many empirical studies, such as Schwartz (1997), Casassus and Collin-Dufresne (2005) and our empirical results in Section 5. It is easy to show that if this proposition is violated an arbitrage will exist.

The bounded futures slope is consistent with market observations of most commodities where deep backwardation commonly exists, but not deep contango. For example, Litzenberger and Rabinowitz (1995) document that the nine-months futures price was strongly backwardated 77% of the time and weakly backwardated 94% of the time between February 1984 and April

1992.

3. Model Specification

In this section, we first emphasize one type of reduced-form model for storable commodities where the convenience yield is modeled by a CIR process. Due to the semi-affine feature of our model, we show that the futures prices can be solved analytically.

3.1 The model

We assume there are three latent factors governing the movements of commodity futures prices which are the stochastic short rate r_t , the convenience yield δ_t and the spot price S_t . To constrain the convenience yield from being negative, we use a Cox–Ingersoll–Ross type specification of the convenience yield δ_t . In the \mathbb{Q} measure, the model is specified as

$$\begin{aligned}
 dr_t &= \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_{1,t}^{\mathbb{Q}}, \\
 d\delta_t &= \kappa_\delta(\theta_\delta - \delta_t)dt + \sigma_\delta\sqrt{\delta_t}dW_{2,t}^{\mathbb{Q}}, \\
 dS_t &= [(r_t - \delta_t)S_t + w]dt + S_t\sigma_{Sr}\sqrt{r_t}dW_{1,t}^{\mathbb{Q}} + S_t\sigma_{S\delta}\sqrt{\delta_t}dW_{2,t}^{\mathbb{Q}} \\
 &\quad + S_t\sqrt{v_0 + v_{S\delta}\delta_t + v_{Sr}r_t}dW_{3,t}^{\mathbb{Q}}, \\
 dW_{1,t}^{\mathbb{Q}}dW_{2,t}^{\mathbb{Q}} &= dW_{3,t}^{\mathbb{Q}}dW_{1,t}^{\mathbb{Q}} = dW_{2,t}^{\mathbb{Q}}dW_{3,t}^{\mathbb{Q}} = 0.
 \end{aligned} \tag{10}$$

W denotes the Wiener process and \mathbb{Q} in the superscript denotes the risk-neutral measure.

Note that the specification of the S_t process indicates a high convenience yield – high spot price volatility relationship, which is consistent with the theory of storage.⁸ In order to make these

⁸ In the theory of storage, the inventory is considered as a buffer to offset demand and supply shocks on commodity markets. High inventory results in low commodity price volatility, and low inventory causes high volatility. Since inventory is negatively correlated with the convenience yield, there exists a high convenience yield – high price volatility relationship. Refer to Pindyck (2001) for details.

factors admissible, we must have $\theta_r > 0, \kappa_r > 0, \theta_\delta > 0, \kappa_\delta > 0$ and $v_0 \geq 0, v_{S\delta} \geq 0, v_{Sr} \geq 0$.

To explain the historical time-series dynamics of prices, we need to specify the risk premia in the relation between the risk-neutral \mathbb{Q} and physical \mathbb{P} measures,

$$d \begin{bmatrix} W_{1,t}^{\mathbb{Q}} \\ W_{2,t}^{\mathbb{Q}} \\ W_{3,t}^{\mathbb{Q}} \end{bmatrix} = \begin{bmatrix} \eta_r \sqrt{r_t} \\ \eta_\delta \sqrt{\delta_t} \\ \eta_s \sqrt{v_0 + v_{S\delta} \delta_t + v_{Sr} r_t} \end{bmatrix} dt + d \begin{bmatrix} W_{1,t}^{\mathbb{P}} \\ W_{2,t}^{\mathbb{P}} \\ W_{3,t}^{\mathbb{P}} \end{bmatrix}. \quad (11)$$

In contrast to previous empirical research (e.g., Schwartz, 1997) that assumes constant risk premia, the specification of risk premia in (11) allows the dependence of the futures risk premia on the convenience yield, which is indicated by Gorton et al. (2007). Also, Bessembinder and Chan (1992) document that the risk premia of futures depends on several state variables. Here, we explicitly assume one of the state variables as the convenience yield since it mainly reflects the demand and supply mismatch and the scarcity of a certain commodity.

Specifically, in the \mathbb{P} measure, the dynamics of the three factors are:

$$\begin{aligned} dr_t &= [k_r(\theta_r - r_t) + \sigma_r \eta_r r_t] dt + \sigma_r \sqrt{r_t} dW_{1,t}^{\mathbb{P}}, \\ d\delta_t &= [k_\delta(\theta_\delta - \delta_t) + \sigma_\delta \eta_\delta \delta_t] dt + \sigma_\delta \sqrt{\delta_t} dW_{2,t}^{\mathbb{P}}, \\ dS_t &= [(r_t - \delta_t + \sigma_{Sr} \eta_r r_t + \sigma_{S\delta} \eta_\delta \delta_t + \eta_s(v_0 + v_{S\delta} \delta_t + v_{Sr} r_t))S_t + w] dt \\ &\quad + S_t \sigma_{Sr} \sqrt{r_t} dW_{1,t}^{\mathbb{P}} + S_t \sigma_{S\delta} \sqrt{\delta_t} dW_{2,t}^{\mathbb{P}} + S_t \sqrt{v_0 + v_{S\delta} \delta_t + v_{Sr} r_t} dW_{3,t}^{\mathbb{P}}, \\ dW_{1,t}^{\mathbb{P}} dW_{2,t}^{\mathbb{P}} &= dW_{1,t}^{\mathbb{P}} dW_{3,t}^{\mathbb{P}} = dW_{2,t}^{\mathbb{P}} dW_{3,t}^{\mathbb{P}} = 0. \end{aligned} \quad (12)$$

Note that this model is not an affine model in the framework of Duffie and Kan (1996), Duffie et al. (2000), and Dai and Singleton (2000). However, although not consistent with reality, if we set $w = 0$, or $w = pS_t$ (for a constant p)⁹ the model transfers to an affine model with three latent factors of r_t ; δ_t and $\ln(S_t)$. Hence, this model has an affine backbone, we thus

⁹ As mentioned before, in (10) we use the storage cost w directly instead of the percentage storage cost p since w is stable cross time but p is not due to the fluctuation of the spot price S_t .

call it a semi-affine model. Besides, under the affine backbone, we also want the model to be maximal in a sense that the maximum number of identifiable parameters exist in the backbone. By following the procedure of Hilliard and Reis (1998), it is easy to show that the model backbone (when $w = 0$) belongs to a "maximal" affine model of $\mathcal{A}_2(3)$ in terms of Dai and Singleton (2000).¹⁰ For brevity, we omit the derivation, the complete details can be provided on request.

3.2 Futures pricing

It is well known (e.g., Cox et al., 1981) that the futures price $F(S, \delta, r, t; T)$ with maturity T at time t follows:

$$F(S, \delta, r, t; T) = E_t^{\mathbb{Q}}[S_T] \quad (13)$$

Since our model is not an affine model, it is not trivial to obtain the futures prices. This is likely to be a reason that no research previously is conducted to model futures with separate storage cost and the convenience yield.

To solve for the futures price F_t , we condition on the history of d_t and r_t , so

$$F_t = E_t^{\mathbb{Q}}[S_T] = E_t^{\delta, r} \left[E_t^S[S_T | \delta_v, r_v, t \leq v \leq T] \right], \quad (14)$$

where the sup- and sub- scripts denote the variable and time that the expectation is taking on.

From Arnold (1974), we know that conditional on the history of δ and r ; S_t in (10) belongs to a special case of general scalar linear equations, hence,

$$E_t^S[S_T | \delta_v, r_v, t \leq v \leq T] = S_T \phi_T + w \int_t^T \frac{\phi_T}{\phi_u} du,$$

where

¹⁰ We also constrain that interest rates from influencing the movement of the convenience yields. This is mainly to make the analytical solution exist.

$$\begin{aligned} \phi_T = \exp \left\{ -\frac{1}{2} v_0 (T - t) - \left(1 + \frac{1}{2} (v_{S\delta} + \sigma_{S\delta}^2) \right) \int_t^T \frac{\phi_T}{\phi_u} du + \left(1 - \frac{1}{2} \sigma_{Sr}^2 \right) \int_t^T r_u du \right. \\ \left. + \int_t^T [\sigma_{Sr} \sqrt{r_t} dW_{1,t}^{\mathbb{Q}} + \sigma_{S\delta} \sqrt{\delta_t} dW_{2,t}^{\mathbb{Q}} + \sqrt{v_0 + v_{S\delta} \delta_t + v_{Sr} r_t} dW_{3,t}^{\mathbb{Q}}] \right\}, \end{aligned}$$

so that the future price is given by

$$F_t = E_t^{\delta,r} \left[E_t^S [S_T | \delta_v, r_v, t \leq v \leq T] \right] = S_t E_t^{\delta,r} [\phi_T] + w \int_t^T E_t^{\delta,r} \left[\frac{\phi_T}{\phi_u} \right] du. \quad (15)$$

Proposition 4. *If the joint process of $\{r_t, \delta_t, S_t\}$ is specified as (10), we have*

$$E_t^{\delta,r} [\phi_T] = e^{\alpha(t,T) + \beta(t,T) + \gamma(t,T)r_t}, \quad (16)$$

$$E_t^{\delta,r} [\phi_T / \phi_u] = e^{A(u;t,T) + B(u;t,T)\delta_t + C(u;t,T)r_t} \quad (17)$$

thus the futures price follows

$$F(S_t, \delta_t, r_t, t; T) = S_t e^{\alpha(t,T) + \beta(t,T) + \gamma(t,T)r_t} + w \times \int_t^T e^{A(u;t,T) + B(u;t,T)\delta_t + C(u;t,T)r_t} du \quad (18)$$

Proof. $E_t^{\delta,r} [\phi_T]$ and $E_t^{\delta,r} [\phi_T / \phi_u]$ can be solved by the Kayman-Kac equation. For the value and detailed derivation of $\alpha(t, T), \beta(t, T), \gamma(t, T), A(u; t, T), B(u; t, T)$, and $C(u; t, T)$, refer to

Appendix E.

Thus, analytical solution for futures prices exists in our model. For further reference, we rewrite (18) as

$$\ln[F(x_t, \delta_t, t; T)] = \alpha(t, T) + x_t + \beta(t, T)\delta_t + \gamma(t, T)r_t + G(t, x_t, \delta_t, r_t; T) \quad (19)$$

with $x_t := \ln(S_t)$

$$\begin{aligned} G(t, x_t, \delta_t, r_t; T) = \ln \left[1 + w e^{-x_t} \int_t^T \exp([A(u; t, T) - \alpha(t, T)] \right. \\ \left. + [B(u; t, T) - \beta(t, T)]\delta_t + [C(u; t, T) - \gamma(t, T)]r_t) du \right] \end{aligned} \quad (20)$$

If $w = 0$, $\ln[F(x_t, \delta_t, t; T)]$ is an affine structure of the three state factors r_t, δ_t and x_t , which is consistent with the model set up when $w = 0$. In other words, the item $G(t, x_t, \delta_t, r_t; T)$ can be

seen as a “correction item” for adding the storage cost w .

Since the interest rates are mainly determined by the treasury bond market, we need to perform a joint estimation of bond and futures prices. The zero-coupon bond price and yield with CIR interest rate process are shown in Brown and Dybvig (1986).

3.3 The probability of negativity in NCY models

In this section, we show that the use of Vasicek or Ornstein–Uhlenbeck (OU) process in modeling the (net) convenience yield encounters a serious problem. Monte Carlo simulations demonstrate that OU processes generate negative convenience yield with high probability (compared with the Vasicek interest-rate model) even in a short period of time such as three months. Models allowing for a negative convenience yield are not arbitrage-free. However, this case contrasts with that of modeling interest rates using the Vasicek model, where negative interest rates are not desirable but cannot result in any arbitrage.¹¹ Therefore, the nonnegative constraint on the convenience yield is not only important theoretically in no-arbitrage pricing theory but also substantial empirically because of the large probability of its violation. In the following, we illustrate the case of modeling NCY as an OU process by comparing it with the Vasicek (1977) interest-rate model, since both of them face problems of non-negative constraints.

Historically, the long-run mean of the interest rate is about 4 or 5% and the volatility is around 1 percent if we estimate a simple Vasicek model. Intuitively, even though such a Vasicek model cannot eliminate the negative interest rate, since the long-term mean is much larger than its volatility, the probability of interest rates going negative is very small, especially in a short

¹¹ It is easy to show that an equivalent martingale measure can be found in the bond market under the Vasicek interest-rate model and hence the Vasicek interest-rate model is arbitrage-free.

period. Hence, the ignorance of the negative interest rate will not have too great an impact on the pricing of interest-rates-related contingent claims, especially when the time to maturity is short. This is not however the case for the convenience yield. Schwartz (1997) estimates the mean and standard deviation of copper's *NCY* in the risk neutral measure as 2.65%¹² and 25.0%, respectively. Given that the copper percentage storage cost is roughly 2%, the probability of the net convenience yield going below the negative percentage storage cost (which corresponds to the negative convenience yield) is similar to that of the hypothetical case where the long-run mean of the convenience yield is 4–5%¹³ but its volatility is as high as 25%. To do a robust check, we follow Bessembinder et al. (1995), Gibson and Schwartz (1990) to obtain the implied convenience yield using one- and three-months copper futures. Then, we perform an AR(1) regression on the implied convenience yield. The mean and volatility of the implied convenience yield are shown in Table 1. Thus, the high volatility of the *NCY* is very likely to make the convenience yield negative even over a short period of time. Apparently, the volatility in the OU process is the key parameter differencing the commodity *NCY* model with the Vasicek interest-rate model. Table 1 compares the parameters in the Vasicek interest-rate model and the commodity *NCY* model.

Next, we explore the probability that the instantaneous convenience yield is below zero in the \mathbb{Q} measure between the current time and an arbitrary future time T . We hence name this probability the probability of negative convenience yield. This probability is also the probability of arbitrage in the *NCY* models. We use parameters in Schwartz (1997) model as an example to calculate the probability of negative convenience yield. We simulate 1000 paths and obtain the

¹² Table 7 of Schwartz (1997) estimates the mean of the convenience yield in the physical measure; we obtain the risk-neutral mean m by deducting the risk premium $\frac{\lambda}{k}$ from the physical measure mean α . Specifically, $\alpha - \frac{\lambda}{k} = 0.248 - \frac{0.256}{1.156} = 0.0265$.

¹³ The distance from the long-run mean of *NCY* (2.25%) to the non-arbitrage barrier (–2%) is 4.25%.

total number of paths that violate the non-negative constraints and then calculate probability of negative convenience yields. Note that, without loss of generality, in this paper we choose copper and oil futures as the embodiments of our study. In the following we list the probability of negative convenience yields with T as 3, 6 and 12 months in Table 2. To compare the probability of negative convenience yield of commodities with the probability of negative interest-rate in the Vasicek model, we also run the same simulation in the Vasicek model. Note that we assume the starting NCY and interest rate to be their long-run mean. The storage costs for oil and copper are shown in Section 4.1.

Table 2 shows that there is a significant chance that arbitrage exists in the NCY models. For example, when using parameters reported in Schwartz (1997) even for $T = 3$ months, the chances are 71% and 73%, respectively of violating no-arbitrage constraints for copper and oil. They are smaller but still significant when calibrating the Schwartz model using data in this paper. On the contrary, it is nearly impossible for interest rates to go negative in the Vasicek model for the same horizon. Since many commodity-related contingent claims are priced by simulating the Schwartz model for a certain period, from several weeks to several years, the existence of carry arbitrage certainly hurts the precision of their pricing. Offering a related example, Section 5 compares values of calendar spread options using models with and without the non-negative convenience yield constraints.

Unlike the traditional approach taken in the literature on modeling the combination of the storage cost and the convenience yield – the net convenience yield, we model them separately in this paper. From (2), it is easy to show that the variance in the net convenience yield is composed of three parts – the variance in the convenience yield, the variance in the inverted spot price and their covariance. This is

$$\text{var}(c_t) = \text{var}(\delta_t) + \text{var}\left(\frac{w}{S_t}\right) - 2\text{cov}\left(\delta_t, \frac{w}{S_t}\right).$$

Since it is commonly known that δ_t and S_t move together (refer to Schwartz, 1997 for example), it is easy to show that $\text{cov}\left(\delta_t, \frac{w}{S_t}\right) \leq 0$, and hence $\text{var}(c_t) \geq \text{var}(\delta_t)$. We thus know that the high degree of convenience yield volatility in *NCY* models is partially caused by the nature of modeling the net convenience yield instead of the convenience yield. Note that as mentioned before, the volatility of the convenience yield is a major factor in determining the probability of the negative convenience yield, so we compare the probability of the negativity in an *NCY* model (for example the Schwartz model) and a model that separately specifies the convenience yield and the storage cost – an alternative to our model with the CIR convenience yield process replaced by the Vasicek process. For details on the alternative model, refer to Appendix F. We calibrate the two models using data from Section 4.1. The probability of the negativity and the volatility of the (net) convenience yield are reported in Table 2. We do see that the volatility of the convenience yield in the Schwartz model is greater than it is in the alternative model, which is very likely causing a higher probability of a negative convenience yield. This confirms our expectation.

Note that although the probabilities of a negative convenience yield in the alternative model is less than in the Schwartz (1997) model, in both cases the probability is significant (more than 10% for 3 months period). Thus, to avoid arbitrage, it is necessary to use a CIR-type convenience yield process.

4. Model calibration – the extended Kalman filter

Since our model is very close to an affine model from both the factor evolution process (12) and the futures prices (19), it is natural to use the Kalman filter to calibrate our model. However, we need to use the extended version of the Kalman filter – the extended Kalman filter.

In the section, we first describe the data, then the empirical methodology and discuss results in the end.

4.1 The data

Our dataset consists of futures contracts on crude oil, copper, and zero-coupon bond prices. For all commodities, we use daily data from January 03 2000 to September 01 2006 (1665 observations for each commodity). The futures prices of *WTI* crude oil (*CL*) and high-grade copper (*HG*) are from the NYMEX.¹⁴ Table 3 contains the summary statistics for commodity prices and returns. The time to maturity ranges from 1 month to 17 months for both oil and copper contracts. We denote F_n as the n th contract closest to maturity; e.g., F_1 is the future contract that is closest to maturity. Since the maturities are in consecutive calendar months, n also roughly denotes the time to maturity (in monthly units). In this paper, we use five time series for oil and copper – F_1, F_5, F_9, F_{13} , and F_{17} contracts. The daily interest rate data of the same period are obtained from the Federal Reserve Board. We use daily 3 and 6 month Constant Maturity Treasury (CMT) yields in our calibration. If we casually compare the F_1 and F_{17} contracts, we see that the difference between the F_1 and F_{17} futures prices alternates signs, which corresponds to periods of backwardation and contango in oil and copper markets.

Lewis (2007) estimate the oil storage cost from 1989 to 2004 is a constant of 0.4 dollar per month per barrel (4.8 dollar per year). We therefore choose \$0.4 per month per barrel as crude oil storage cost. For storage cost information on industrial metals, we went through various issues of LME public announcements and publications on warehouse rents. Fig. 1 shows historical copper storage cost on NYMEX, where the storage cost changes quite slowly and

¹⁴ Note that the oil and copper data are from the NYMEX and COMEX divisions, respectively.

predictably comparing the large volatility of the copper price. We choose average storage cost of 4.75 per month per short ton for copper.

To show that the storage cost is relatively stable over a certain period, Table 4 lists current storage costs for various commodities traded in LME, which are constants for one year period.

4.2 The extended Kalman filter

One of the difficulties in the calibration of our model is that the three factors are not directly observable. Several calibration methodologies have been proposed to solve this problem, such as efficient method of moments (Gallant and Tauchen, 1996), maximum likelihood estimation (i.e., the Chen and Scott, 1993 method), and the Kalman filter method. Duffee and Stanton (2004) compare these methods and conclude that the (extended) Kalman filter is the best method among those three, especially when the model is complicated. Maximum likelihood estimation produces strongly biased parameters when modeling complex dynamics and the efficient method of moments is not even acceptable. Therefore in this paper we use the extended Kalman filter to calibrate our model, however since our model is “semi-affine”, we need to use the extended Kalman filter, which is capable of solving non-linear filtering problems especially when the non-linearity is not very high.¹⁵ Many previous studies have used the Kalman filter to estimate the CIR-type latent factors, especially in interest-rate modeling, such as Duffee and Stanton (2004), Chen and Scott (2003).

Note that in a system of n commodities (n is generally very large), models of all commodities should be calibrated together since all commodities share the same interest rates

¹⁵ There is an excellent summary of extended Kalman filter on non-linear statespace models written by Orderud, F., which can be accessed on line <http://www.idi.ntnu.no/~fredrior/files/orderud05sims.pdf>

process. This will, however, involve an intensive computational load and hence is not a plausible strategy in reality. We thus follow Casassus and Collin-Dufresne (2005) to estimate individual commodities. Under this procedure, however, interest rates backed out from various commodities may behave differently.¹⁶

The state-space form normally consists of a transition equation and a measurement equation. The transition equation shows the stochastic process involved in generating data. Thus, the transition equation in the model should be the discrete version of (12) with $x_t := \ln S_t$. The stochastic process of x factor is

$$\begin{aligned}
 dx_t &= \frac{1}{S_t} dS_t - \frac{1}{2} \frac{1}{S_t^2} (dS_t)^2 \\
 &= [(r_t - \delta_t + \sigma_{Sr} \eta_\delta \delta_t + \eta_S (v_0 + v_{S\delta} \delta_t + v_{Sr} r_t)) \\
 &\quad + w e^{-x_t} - \frac{1}{2} (\sigma_{Sr}^2 r_t + \sigma_{S\delta}^2 \delta_t + v_0 + v_{S\delta} \delta_t + v_{Sr} r_t)] dt \\
 &\quad + \sigma_{Sr} \times \sqrt{r_t} dW_{1,t}^{\mathbb{P}} + \sigma_{S\delta} \sqrt{\delta_t} dW_{2,t}^{\mathbb{P}} \\
 &\quad + \sqrt{v_0 + v_{S\delta} \delta_t + v_{Sr} r_t} dW_{3,t}^{\mathbb{P}}
 \end{aligned} \tag{21}$$

The measurement equation relates the time series for the multivariate observable variables (futures and bonds prices for various maturities in our case) to an unobservable vector of state variables δ, r, x . The measurement equation is obtained using (19) with uncorrected noises taking account of the pricing errors. These errors may be caused by bid-ask spreads, non-simultaneity of the observations, etc. In the following we state the transition and measurement equations in more details.

Let $\Delta = t_{n+1} - t_n$ be the interval between two observations and $X = (X_1, \dots, X_T)$ and $Z = (Z_1, \dots, Z_T)$ be the total latent state variables and observations (from 1 to T). $X_n =$

¹⁶ We thank the referee for pointing this out.

$[r_n, \delta_n, x_n]$ represents the vector for state variables at time t_n , and

$$Z_n = [\ln(F(t, T_1)), \dots, \ln(F(t, T_5))],$$

$y(t, T_{r1}), \dots, y(t, T_{r2})]$ represents 7 observations of each time with the first five as the futures prices (F_1, F_5, F_9, F_{13} and F_{17}) and the last two as bond yields (3 and 6 months).

Since the data used is in a daily frequency, which is quite short, the Euler discretization of (12) can serve as a good proxy for the transition equation,¹⁷ i.e.

$$X_n = f(X_{n-1}) + \omega_n, \quad (22)$$

where

$$f(X_{n-1}) = MX_{n-1} + U(X_{n-1}) + L$$

M, L are constants, $U(X_{n-1})$ and is a vector containing the non-linear part in the equation. Thus,

$$M = \begin{bmatrix} 1 + (\sigma_r \eta_r - k_r) \Delta & 0 & 0 \\ 0 & 1 + (\sigma_\delta \eta_\delta - k_\delta) \Delta & 0 \\ \left(1 + \sigma_{Sr} \eta_r + \eta_S v_{Sr}\right) \Delta & \left(\sigma_{S\delta} \eta_\delta + v_{S\delta} \eta_S\right) \Delta & 1 \end{bmatrix},$$

$$L = \left[k_r \theta_r \quad k_\delta \theta_\delta \quad v_0 \eta_S - \frac{1}{2} v_0 \right]^T \Delta, U(X_{n-1}) = [0 \quad 0 \quad w e^{-x} \Delta]^T.$$

ω_n is a random noise with zero mean and variance

$$Q_n = \begin{bmatrix} \sigma_r^2 r_{n-1} & 0 & \sigma_{Sr} \sigma_r r_{n-1} \\ 0 & \sigma_r^2 \delta_{n-1} & \sigma_{S\delta} \sigma_\delta \delta_{n-1} \\ \sigma_{Sr} \sigma_r r_{n-1} & \sigma_{S\delta} \sigma_\delta \delta_{n-1} & v_0 + v_{S\delta} \delta_{n-1} + v_{Sr} r_{n-1} + \sigma_{Sr}^2 r_{n-1} + \sigma_{S\delta}^2 \delta_{n-1} \end{bmatrix} \Delta.$$

From the futures pricing formula, we get the measurement equation,

$$Z_n = h(X_n) + \varepsilon_n, \quad (23)$$

where

$$h(X_n) = V_n X_n + D_n + P(X_n)$$

¹⁷ If the sample interval is relatively longer, such as the monthly frequency in the fixed income literature (for example Duffee, 1999), simple Euler discretization might be too rough, hence, the exact solution of $X_n | X_{n-1}$ should be calculated based on the continuous stochastic process.

with

$$V_n = \begin{bmatrix} \gamma(t, T_1) & \beta(t, T_1) & 1 \\ \vdots & \vdots & \vdots \\ \gamma(t, T_5) & \beta(t, T_5) & 1 \\ \zeta(t, T_{r1}) & 0 & 0 \\ \zeta(t, T_{r2}) & 0 & 0 \end{bmatrix} \quad (24)$$

$$D_n = [\alpha(t, T_1) \dots \alpha(t, T_5) \ \pi(t, T_{r1}) \ \pi(t, T_{r2})]^T, \quad (25)$$

$$P(X_n) = [G(t_n, x_t, \delta_n, r_n; T_1) \dots G(t_n, x_t, \delta_n, r_n; T_5) \ 0 \ 0]^T, \quad (26)$$

$P(X_n)$ contains the nonlinear part of the measurement equation. We assume ε_n follows a iid joint normal distribution with zero means and a diagonal variance covariance matrix R

$$R := \text{diag}(\xi_{F1}^2, \dots, \xi_{F5}^2, \xi_{r1}^2, \xi_{r2}^2).$$

4.3 The results

In this section, we discuss primarily the model calibration results for oil and copper. Table 5 presents the estimates of our model. For each commodity, we present the risk-neutral and the risk premia parameters. Table 5 shows that nearly all parameters are significant. This implies that our model setup is indeed necessary to explain the dynamics of the two commodities. The positive sign of η_δ and η_s shows that both the convenience yield and spot price have positive risk premia. The risk premia also depend on the latent variables such as the convenience yield and the interest rates. We define the steady-state risk premia – the risk premia when the latent variables are in their steady states (or the long-run means)– for the convenience yield λ_δ and spot price λ_s as

$$\lambda_\delta = \eta_\delta \sqrt{\theta_\delta}, \quad (27)$$

$$\lambda_s = \eta_s \sqrt{v_0 + v_{s\delta} \theta_\delta + v_{sr} \theta_r}.$$

Moreover, $\sigma_{s\delta}$ is positive and significant for both oil and copper in that an increase in the

convenience yield is correlated with an increase of spot price, which is consistent with the theory of storage. In the mean while, it also indicates the high convenience yield – high spot price volatility relationship. We also see that the risk premium for interest rate r is negative. This is easy to understand. Bond prices should have positive risk premium since they are tradable assets; however, because interest rates are negatively related to bond prices, interest rates should thus have negative risk premia (see Duffee, 2002 for a detailed discussion).

For oil, the convenience yield process is estimated to be persistent under both the risk-neutral and the physical measure with a long-run mean of about 22% and a half-life of about 1.5 years. The steady-state risk premia for the convenience yield and spot prices are $\lambda_\delta = 0.190$, and $\lambda_S = 0.624$, respectively. From the small pricing error ξ_{F1} to ξ_{F5} , we can see that our model fits the futures prices quite well. The average error is around 0.0012, the average absolute deviation is 0.0203, which is around 0.6% of the log price of the nearby futures.

For copper, the convenience yield process is persistent under both the risk-neutral and the physical measures with a long-run mean of about 9% and a half-life of about 2 year. The steady-state risk premia for the convenience yield and spot prices are $\lambda_\delta = 0.212$, and $\lambda_S = 1.097$, which are larger than those for oil. From the small pricing error ξ_{F1} to ξ_{F5} , we can see that our model fits the copper futures prices quite well. The average error is quite small, around 0.0010, the average absolute deviation is 0.014, which is around 0.3% of the log price of the nearby futures.

5. Application –Calendar Spread Options

A calendar spread option is an option contract based on the spread between two maturities of the same commodity. The calendar futures spread is associated with inventory management of a certain commodity. Assume at time t , the first futures is $F(t, T_1)$ with maturity

T_1 , the second futures is $F(t, T_2)$ with maturity T_2 , ($T_1 \leq T_2$), the locked profit (or loss) of holding one unit of inventory (or strategic storage) between T_1 and T_2 is $F(t, T_2) - F(t, T_1) - w(T_2 - T_1)$. Since as mentioned before the storage cost $w(T_2 - T_1)$ can be considered as a constant, thus, the profit or loss of holding inventory from T_1 to T_2 mainly depends on the calendar spread of the futures prices at time t . Also, many firms own an *real* calendar spread option with the strike as storage cost ($w(T_2 - T_1)$) because they can choose not to store inventory when the calendar spread is very small. In the meanwhile, as more and more investors choose to invest commodity index, options on calendar spreads will enable customers to mitigate the risk of changing price differentials between successive futures contract months when rolling from one futures contract month to another. Introduced in 2001 on NYMEX, the calendar spread options of oil products have become very popular as a hedging tool for oil refinery firms. Currently, calendar spread options have been traded for many commodities, such as gasoline, heating oil, natural gas, wheat, corn, wheat, soybeans, soybean meal and soybean oil.¹⁸

The spread call and put option prices are denoted as c and p at current time t_0 with maturity as T_0 ($T_0 \leq T_1 \leq T_2$). Thus, the payoff of the call and put calendar spread options at M is $[F(T_0, T_1) - F(T_0, T_2) - K, 0]^+$ and $[K - (F(T_0, T_1) - F(T_0, T_2)), 0]^+$, respectively. In this paper, we only focus on the pricing of put options for briefness, calls can be calculated following the same principle and thus are omitted here. The put calendar spread option is priced as

$$p = E_{t_0}^{\mathbb{Q}} \left[\exp \left(\int_{t_0}^T r_u du \right) [K - F(T_0, T_1) + F(T_0, T_2), 0]^+ \right] \quad (28)$$

For the best of the author's knowledge, the analytical solution for calendar spread options is not available if $K \neq 0$. Thus, to price the calendar spread option, we use Monte Carlo simulation. In

¹⁸ The calendar spread option for agricultural products can also help growers manage the risk of the differential of the old crop and new crop months.

this section, we simulate the spread option using two models – our model and the Schwartz three-factor model (1997). Note that the futures process for the Schwartz (1997) three-factor model is derived in Hilliard and Reis (1998), where the futures process does not depend on any of the state variables. In our model, the process for futures $F(t, T)$ in the risk-neutral measure can be derived from (19)

$$\begin{aligned} \frac{dF(t, T)}{F(t, T)} = & \left[\left(\gamma(t, T) + \frac{\partial G(r, x, \delta, t; T)}{\partial r} \right) \sigma_r + \sigma_{Sr} \right] \sqrt{r_t} dW_{1,t}^{\mathbb{Q}} \\ & + \left[\left(\beta(t, T) + \frac{\partial G(r, x, \delta, t; T)}{\partial \delta} \right) \sigma_\delta + \sigma_{S\delta} \right] \sqrt{\delta_t} dW_{2,t}^{\mathbb{Q}} + \sqrt{v_0 + v_{S\delta}\delta_t + v_{Sr}r_t} dW_{3,t}^{\mathbb{Q}}, \end{aligned} \quad (29)$$

Note that the futures price, which is a martingale, depends on the state variables in our model. Thus, we need to simulate state variables together with the futures prices in our simulation.

We choose the crude oil futures contracts for example with $T_1 = 13$ months, $T_2 = 17$ months, $T_0 = 13$ months.¹⁹ For convenience we define $\zeta(t) := F(t, T_1) - F(t, T_2)$ as the futures spreads and the initial spread as $\zeta(t_0)$.²⁰ According to the initial oil prices, we also adjust the interval of the strike prices for different pricing days.

In order to make the simulation accurate, we use anti-variate techniques in generating random variables. Moreover, in order to make the two models more comparable to each other, we use the same set of random numbers for both models. We select four days to calculate the option value – the first (January 03, 2000), 500th, 1000th and last day (September 01, 2006) of our sample. We can see these four days cover different convenience yield scenarios ($\delta_t = 0.53, 0.27, 0.31, \text{ and } 0.04$, respectively). We simulate 1000 paths for both models and Table 6 shows the results.

¹⁹ It is a convention that $T_1 = T_0$ for the calendar spread option on NYMEX.

²⁰ It is a convention that the spread is defined as the shorter-term futures minus the longer term futures on NYMEX. But for convenience of presentation, we also define the other way of constructing spreads.

Depending on the different state factors, our model produces different option prices relative to Schwartz model. However, the prices of out-of-money put in our model is always lower than that in the Schwartz model for all days. This is mainly because of the constraint of non-negativity of the convenience yield in our model. Specifically, at time T_1 , as we know the minimal spread $\underline{\zeta}(T_1)$ occurs when the convenience yield δ_t is always zero between time T_1 and T_2 (i.e. $\delta_t = 0, T_1 \leq t \leq T_2$). In this case $\underline{\zeta}(T_1)$ can be calculated as

$$\underline{\zeta}(T_1) = \inf_{\delta} \{F(T_1, T_1) - F(T_1, T_2)\} = -E_{T_1}^{\mathbb{Q}} \left[\int_{T_1}^{T_2} S_u r_u du \right] - w(T_2 - T_1). \quad (30)$$

Assuming independency between spot price and interest rates, (30) changes to

$$\underline{\zeta}(T_1) = - \int_{T_1}^{T_2} E_{T_1}^{\mathbb{Q}}[S_u] E_t^{\mathbb{Q}}[r_u] du - w(T_2 - T_1) \quad (31)$$

$$= - \int_{T_1}^{T_2} F(T_1, u) f(T_1, u) du - w(T_2 - T_1) \quad (32)$$

Thus, in our model the maximal payoff po_{T_1} for a put option at T_1 is

$$\sup_{\delta} \{po_{T_1}\} = \sup_{\delta} \{[K - \zeta(T_1), 0]^+\} = \left[\sup_{\delta} \{K - \zeta(T_1)\}, 0 \right]^+ = [K - \underline{\zeta}(T_1), 0]^+$$

Thus, in our model there is an upper boundary that the payoff of the spread put option cannot exceed. However, *NCY* models do not place non-negative constraints on the convenience yield, and thus do not have the upper bound of the put option. Hence our model generates lower prices for spread options for out-of-money puts than the Schwartz (1997) model. In some extreme cases, when $K \leq \underline{\zeta}(T_1)$, the maximal payoff for the put option is zero; hence the put option value for that strike must be zero. (Refer to the result for the last day in Table 5.) Similar arguments apply to in-the-money call spread options.

6. Conclusion

This paper investigates primarily arbitrage-free conditions under the dual cash and storage markets for commodities. We prove that if there is no arbitrage existing in the cash-storage dual markets, the convenience yield must be non-negative. However, as we show in this paper, the classical *NCY* models do exhibit a high probability of violating the non-negativity criteria of the convenience yield due mainly to the high degree of volatility of the convenience yield. Thus these models are not arbitrage free.

In contrast to existing *NCY* models, we propose a three-factor model (capturing log-spot price, the convenience yield and the interest rate), in which the convenience yield and the storage cost are treated separately. We assume the convenience yield following a CIR-type process to assure its non-negativity, with the storage cost treated as a constant. We illustrate that the separation of the convenience yield and the storage cost reduces the volatility of the convenience yield, and hence yields a smaller probability of violating the non-negativity criteria. More importantly, in our model futures prices have an analytical solution, which makes it attractive from a practical standpoint. We use the extended Kalman filter to estimate our model and use oil and copper as two examples for model calibration. We show that, because of the non-negativity of the convenience yield, the slope of the futures term structure cannot exceed a certain threshold, and our model has a lower price for deep out-of-the-money put calendar spread options relative to the Schwartz (1997) model.

Acknowledgements

We thank Henry Cao, Jaime Casassus, Pierre Collin-Dufresne, Jun Liu, Yaping Wang, Wei Xiong, seminar participants at the Five Star Conference in Beijing for help comments. We are also grateful to an anonymous referee whose suggestions greatly improved the paper.

Appendix A. Proof of Proposition 1

Under the \mathbb{Q} measure the futures prices F_t is expressed as,

$$F(t, s) = E_t^{\mathbb{Q}}[F(s, s)] = E_t^{\mathbb{Q}}[S_s]. \quad (33)$$

Thus, denoting \mathcal{F}_t as filtration up to time t , we thus have

$$E_t[dS_t | \mathcal{F}_t] = \lim_{\Delta \rightarrow 0} E_t^{\mathbb{Q}}[S_{t+\Delta} | \mathcal{F}_t] - S_t = \lim_{\Delta \rightarrow 0} F(t, t + \Delta) - F(t, t) = \frac{\partial F(t, s)}{\partial s} \Big|_{s=t} dt \quad (34)$$

From (1), we thus obtain

$$E_t^{\mathbb{Q}}[dS_t] = E_t^{\mathbb{Q}}[(S_t(r_t - \delta_t) + w)dt | \mathcal{F}_t] = (S_t(r_t - \delta_t) + w)dt \quad (35)$$

Thus, if (1) holds, (35) holds.

In the meanwhile, if (35) holds

$$dS_t = (S_t(r_t - \delta_t) + w)dt + v_t d\beta_t^{\mathbb{Q}},$$

where $\beta_t^{\mathbb{Q}}$ is a $n \times 1$ vector of Brownian motion and v_t is a $n \times 1$ vector adapted to the filtration

\mathcal{F}_t . Hence,

$$S_s = S_t + \int_t^s (S_u(r_u - \delta_u) + w)du + \int_t^s v_u d\beta_u^{\mathbb{Q}}$$

and hence

$$F(t, s) = E_t^{\mathbb{Q}}[S_s] = S_t + \int_t^s (S_u(r_u - \delta_u) + w)du.$$

Thus, if and only if (1) holds the drift of the spot price follows $E_t^{\mathbb{Q}}[dS_t] = (S_t(r_t - \delta_t) + w)dt$.

In the next part, we show that (1) and (6) are equivalent. Since (1) and (35) are equivalent if we can prove that (35) and (6) are equivalent as well, then we prove that (1) and (6) are equivalent. In the following we show that (35) and (6) are equivalent. Assume that

$$dS_t = \iota_t dt + v_t d\beta_t^{\mathbb{Q}}.$$

From Feynman–Kac formula (refer to Oksendal, 2005), we know that if (6) holds, then

$$\iota_t = -\left(\delta_t - \frac{w}{S_t} - r_t\right)S_t = (r_t - \delta_t)S_t + w \quad (36)$$

also if (36) holds, Feynman–Kac shows that S_t must satisfy (6). Hence, (35) and (6) are equivalent, and hence (1) and (35) are equivalent.

Appendix B. Proof of Lemma 1.

For any self-financing trading strategy ϕ in the \mathbb{Q} measure, wealth process $V(\phi)$ equals

$$\begin{aligned} V_t(\phi) &= \phi_t^B \cdot B_t + \phi_t^F \cdot F_t + \phi_t^S \hat{S}_t + \phi_t^M M_t \\ &= \phi_t^B \cdot B_t + \phi_t^S S_t + \phi_t^M M_t, \end{aligned} \quad (37)$$

where $\phi_t^B, \phi_t^F, \phi_t^S$ and ϕ_t^M are portions of bonds, futures, spot commodity and money market

account at time t . The diffusion parameters for bonds, futures and spot are respectively

$\sigma_B, \sigma_F, \sigma_S$, and $\beta_t^{\mathbb{Q}}$ denotes an n -dimensional Brownian motion. We also assume the initial wealth $V_0(\phi) = 0$.

Note that to enter into a futures prices does not need any initial payment, thus $F_t = 0$.

$V(\phi)$ satisfy the following condition

$$V_t(\phi) = \int_0^t \phi_u^B \cdot dB_u + \int_0^t \phi_u^F \cdot dF_u + \int_0^t \phi_u^S \cdot d\hat{S}_u + \int_0^t \phi_u^M \cdot dM_u.$$

Note that the holder of physical commodity has to pay out the storage cost. The discounted wealth process $V_t^*(\phi)$ in \mathbb{Q} measure satisfy

$$dV_t^*(\phi) = \frac{\phi_t^F}{M_t} \cdot dF_t + \frac{\phi_t^B}{M_t} \cdot (dB_t - r_t B_t dt) + \frac{\phi_t^S}{M_t} d\hat{S}_t - \frac{r_t}{M_t} \phi_t^S S_t dt. \quad (38)$$

From (3) and (5),

$$\begin{aligned} dV_t^*(\phi) &= \frac{\phi_t^S}{M_t} d\hat{S}_t - \frac{r_t}{M_t} \phi_t^S S_t dt \\ &= \frac{S_t}{M_t} (-\phi_t^S \delta_t) dt + \frac{1}{M_t} (\phi_t^F \cdot \sigma_F + \phi_t^B \cdot \sigma_B + \phi_t^S \sigma_S) d\beta_t^{\mathbb{Q}} \end{aligned} \quad (39)$$

If $\delta_t = 0$, the wealth process ϕ is a martingale for any trading strategy, hence \mathbb{Q} is a martingale measure for cash-storage dual markets. Thus, arbitrage does not exist in the cash-storage market.

However, for any time t when $\delta_t \neq 0$, we can always find a proper trading strategy

$\tilde{\phi}_t := [\tilde{\phi}_t^B \tilde{\phi}_t^F \tilde{\phi}_t^S]$ to make $\tilde{\phi}_t^F \cdot \sigma_F + \tilde{\phi}_t^B \cdot \sigma_B + \tilde{\phi}_t^S \sigma_S = 0$ and $\tilde{\phi}_t^S \delta_t < 0$. Thus, in this case

$$V_t^*(\tilde{\phi}) = \int_0^T \tilde{\phi}_t V_t^*(\tilde{\phi}) = \int_0^T \frac{S_t}{M_t} (-\tilde{\phi}_t^S \delta_t) dt > 0.$$

Following the numeraire invariance theorem (Duffie, 2006), the strategy $\tilde{\phi}$ make $V_T(\tilde{\phi})$ amount of money with zero investment at time 0. Therefore, if no arbitrage exists in the cash-storage dual markets, $\delta_t \equiv 0$ must hold.

Appendix C.

Arbitrage scenario, when $\delta_t > 0$ and short-sell of physical commodity is allowed

Before stating the arbitrage strategy, we first address the mechanism of the “short-selling” in the storage market. Assume an agent A_S wants to lend 1 unit physical commodity to an investor I_S at t (when the spot price is S_t), he should only receive $\left(1 - \frac{w\Delta t}{S_{t+\Delta t}}\right)$ amount of commodity at time $t + \Delta t$ because it makes no difference for I_S by storing his/her commodity or lending it out.

If the convenience yield δ_t is positive for a short interval Δt ; I_S short-sells 1 unit of physical commodity and get S_t amount of money and save it in the bank, in the mean time he/she buys one futures matured at $t + \Delta t$.

At time $t + \Delta t$, I_S pays $F(t, t + \Delta t)$ to get the 1 unit physical commodity from the futures position and return $\left(1 - \frac{w\Delta t}{S_{t+\Delta t}}\right)$ amount physical commodity to the lending agent.²¹ Therefore, his payoff $po_{t+\Delta t}$ at $t + \Delta t$ is,

$$po_{t+\Delta t} = S_t + S_t r_t \Delta t + \frac{w\Delta t}{S_{t+\Delta t}} \times S_{t+\Delta t} - F(t, t + \Delta t) = S_t \delta_t \Delta t > 0 \quad (40)$$

Thus, I_S has a zero investment at time t but makes a positive payoff at time $t + \Delta t$. Thus, the above strategy offers an arbitrage for I_S when $\delta_t > 0$. If δ_t is positive for a horizon H from t to $t + H$, I_S can continuously conduct this strategy and hence his arbitrage payoff at $t + H$ would be

$$po_{t+H} = \lim_{\Delta t \rightarrow 0} \sum_j \left(S_{t_j} \delta_{t_j} \Delta t \right) = \int_t^{t+H} S_s \delta_s ds > 0.$$

²¹ Since Δt is very short, we assume there is a mark-to-market at t and $t + \Delta t$, but not in between.

Appendix D. Proof of Theorem 1

From the proof of Lemma 1, for any self-financing trading strategy ϕ with zero initial wealth $V_0(\phi) = 0$, the deflated wealth process in the \mathbb{Q} measure at time T , $V_T^*(\phi)$ is

$$V_T^*(\phi) = \int_0^T \frac{S_t}{M_t} (-\phi_t^S \delta_t) dt + \int_0^T \times \frac{1}{M_t} (\phi_t^F \cdot \sigma_F + \phi_t^B \cdot \sigma_B + \phi_t^S \cdot \sigma_S) d\beta_t^{\mathbb{Q}}. \quad (41)$$

If $\delta_t < 0$, we can find a trading strategy $\tilde{\phi}_t = [\tilde{\phi}_t^B \tilde{\phi}_t^F \tilde{\phi}_t^S]$ with $\tilde{\phi}_t^S > 0$, which also satisfies

$\tilde{\phi}_t^F \cdot \sigma_F + \tilde{\phi}_t^B \cdot \sigma_B + \tilde{\phi}_t^S \sigma_S = 0$. Thus, at time T

$$V_T^*(\tilde{\phi}) = \int_0^T \frac{S_t}{M_t} (-\tilde{\phi}_t^S \delta_t) dt > 0$$

and thus an arbitrage exists for the trading strategy $\tilde{\phi}$. If $\delta_t \geq 0$, since $\phi_t^S \geq 0$, $V_T^*(\phi)$ is a supermartingale,

$$E_0^{\mathbb{Q}}[V_T^*(\phi)] \leq V_0^*(\phi) = 0.$$

Thus, an arbitrage strategy such that $V_0^*(\phi) = 0$, and $V_T^*(\phi) \geq 0$, and $P(V_T^*(\phi) > 0) > 0$ can not exist.

Appendix E. Proof of Proposition 4

The future prices need to satisfy the following equation,

$$\begin{aligned}
 0 = & \frac{\partial F}{\partial S} [(r - \delta)S + w] + \frac{\partial F}{\partial \delta} k_\delta (\theta_\delta - \delta) + \frac{\partial F}{\partial r} k_r (\theta_r - r) + \frac{1}{2} \times \frac{\partial^2 F}{\partial r^2} \sigma_r^2 r \\
 & + \frac{1}{2} \frac{\partial^2 F}{\partial \delta^2} \sigma_\delta^2 \delta + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} S^2 [(v_{x\delta} + \sigma_{S\delta}^2) \delta + (\sigma_{Sr}^2 + v_{xr}) r + v_0] \\
 & + \frac{\partial^2 F}{\partial S \partial \delta} \sigma_\delta \sigma_{S\delta} \delta S + \frac{\partial^2 F}{\partial S \partial r} \sigma_r \sigma_{Sr} S r + \frac{\partial F}{\partial t},
 \end{aligned} \tag{42}$$

with boundary condition $F(S, \delta, r, T; T) = S$.

Next, we solve $E_t^{\delta, r}[\phi_T]$. Define $V_t := S_t E_t^{\delta, r}[\phi_T]$, hence V_t should satisfy the PDE of (42) with $w = 0$, i.e. if we set $w = 0$. Observing that the coefficients of the PDE are affine in δ and r , we postulate a solution of the form

$$V(x, \delta, r, t; T) = S \exp\{\alpha(t; T) + \beta(t; T) + \gamma(t; T)r\},$$

where α, β and γ are some arbitrary functions to be determined. Substitution into the PDE then yields

$$\begin{aligned}
 & r - \delta + \beta k_\delta (\theta_\delta - \delta) + \gamma k_r (\theta_r - r) + \frac{1}{2} \sigma_\delta^2 \delta \beta^2 + \frac{1}{2} \sigma_r^2 \gamma^2 r + \beta \sigma_\delta \sigma_{S\delta} \delta \\
 & + \gamma \sigma_r \sigma_{Sr} r + \frac{\partial \alpha}{\partial t} + \delta \frac{\partial \beta}{\partial t} + r \frac{\partial \gamma}{\partial t} = 0.
 \end{aligned}$$

Note that this PDE contains exclusively of terms that are independent of δ and r and terms that are proportional to δ . But since this PDE must hold for all values of δ and r , it must be the case that terms independent of δ and r are equal to zero, and terms proportional to δ and r are also equal to zero, for all δ and r . Therefore, we get three ODEs

$$\frac{\partial \beta}{\partial t} + (\sigma_\delta \sigma_{S\delta} - k_\delta) \beta + \frac{1}{2} \sigma_\delta^2 \beta^2 - 1 = 0, \tag{43}$$

$$\frac{\partial \gamma}{\partial t} + (\sigma_r \sigma_{sr} - k_r) \gamma + \frac{1}{2} \sigma_r^2 \gamma^2 + 1 = 0, \quad (44)$$

$$\beta k_\delta \theta_\delta + \gamma k_r \theta_r + \frac{\partial \alpha}{\partial t} = 0, \quad (43)$$

with terminal conditions $\alpha(T) = \beta(T) = \gamma(T) = 0$.

The first ODE is a Riccati equation, which can be solved by the substitution $\beta = \frac{2u'}{\sigma_\delta^2 u}$.

This gives

$$Tu'' + 2au' - bu = 0,$$

where $a = \sigma_\delta \sigma_{s\delta} - k_\delta$ and $b = \sigma_\delta^2$. The solution for u is therefore $u(t) = K_1 e^{\lambda_1 t} + K_2 e^{\lambda_2 t}$,

where

$$\lambda_1 = -\frac{1}{2}a - \frac{1}{2}\sqrt{a^2 + 2b}, \quad \lambda_2 = -\frac{1}{2}a + \frac{1}{2}\sqrt{a^2 + 2b}$$

so that

$$\beta(t) = \frac{2\lambda_1 e^{\lambda_1 t} + 2K\lambda_2 e^{\lambda_2 t}}{\sigma_\delta^2 (e^{\lambda_1 t} + K e^{\lambda_2 t})},$$

where $K = K_2/K_1$. The terminal condition $\beta(T) = 0$ implies that

$$K = -\frac{\lambda_1}{\lambda_2} e^{(\lambda_1 - \lambda_2)T}.$$

so we have

$$\beta(t, T) = \frac{2\lambda_1}{\sigma_\delta^2} \left[\frac{1 - e^{(\lambda_1 - \lambda_2)(T-t)}}{1 - \frac{\lambda_1}{\lambda_2} e^{(\lambda_1 - \lambda_2)(T-t)}} \right]. \quad (46)$$

Similarly, for γ , we assume $\gamma = \frac{2u'}{\sigma_r^2 u}$. This gives

$$2u'' + 2cu' + du = 0,$$

where $c = \sigma_r \sigma_{sr} - k_r$ and $d = \sigma_r^2$. The solution for u is therefore

$$u(t) = K_3 e^{\lambda_3 t} + K_4 e^{\lambda_4 t},$$

where

$$\lambda_3 = -\frac{1}{2}c - \frac{1}{2}\sqrt{c^2 - 2d}, \quad \lambda_4 = -\frac{1}{2}c + \frac{1}{2}\sqrt{c^2 - 2d}$$

here we make a technical assumption that $c^2 \geq 2d$, by preliminary studies we know that k_r is in a magnitude of 0.1, σ_r is about 0.04, and the correlation of spot and interest rates is typically close to zero (refer to Schwartz, 1997). If $c^2 < 2d$, we refer readers to Birkhoff and Gian-Carlo (1989) and Kim and Omberg (1996). Thus,

$$\gamma(t, T) = \frac{2\lambda_3}{\sigma_r^2} \left[\frac{1 - e^{(\lambda_3 - \lambda_4)(T-t)}}{1 - \frac{\lambda_3}{\lambda_4} e^{(\lambda_3 - \lambda_4)(T-t)}} \right]. \quad (47)$$

We use this to solve the second ODE to get α

$$\begin{aligned} \alpha(t, T) &= k_\delta \theta_\delta \int_t^T \beta(u) du + k_r \theta_r \int_t^T \gamma(u) du \\ &= \frac{2k_\delta \theta_\delta}{\sigma_r^2} \left[(T-t)\lambda_1 + \log \left(1 - \frac{\lambda_1}{\lambda_2} \right) - \log \left(1 - \frac{\lambda_1}{\lambda_2} e^{(\lambda_1 - \lambda_2)(T-t)} \right) \right] \\ &\quad + \frac{2k_r \theta_r}{\sigma_r^2} \left[(T-t)\lambda_3 + \log \left(1 - \frac{\lambda_3}{\lambda_4} \right) - \log \left(1 - \frac{\lambda_3}{\lambda_4} e^{(\lambda_3 - \lambda_4)(T-t)} \right) \right] \end{aligned} \quad (48)$$

Thus,

$$E_t^{\delta, r}[\phi_T] = \frac{V_t}{S_t} = \exp\{\alpha(t; T) + \beta(t; T)\delta + \gamma(t; T)r\}.$$

To solve for the futures price $F(x, \delta, r, t; T)$ when $w > 0$ we need to evaluate the term

$\int_t^T E_t^{\delta, r}[\phi_T / \phi_u] du$. The expectation can be evaluated by conditioning on δ_u and r_u , define

$H(\delta, r, t; u, T) := E_t^{\delta, r}[\phi_T / \phi_u]$, from the iterated expectation theorem we have

$$H(\delta, r, t; u, T) := E_t^{\delta, r} \left[E_t^{\delta, r} \left[\frac{\phi_T}{\phi_u} \right] \right] = E_t^{\delta, r} \left[e^{\alpha(u, T) + \delta_u \beta(u, T) + r_u \gamma(u, T)} \right].$$

$H(\delta, r, t; u, T)$ satisfy

$$\frac{\partial H}{\partial \delta} k_\delta (\theta_\delta - \delta) + \frac{1}{2} \frac{\partial^2 H}{\partial \delta^2} \sigma_\delta^2 \delta + \frac{\partial H}{\partial r} k_r (\theta_r - r) + \frac{1}{2} \frac{\partial^2 H}{\partial r^2} \sigma_r^2 r + \frac{\partial H}{\partial t} = 0, \quad (49)$$

with boundary condition $H(\delta, r, t; u, T) = e^{\alpha(u, T) + \delta u \beta(u, T) + r u \gamma(u, T)}$. Assume

$H = \exp(A(t; u, T) + \delta_t B(t; u, T) + r_t C(t; u, T))$, A , B , and C satisfy respectively, the

following equations:

$$\frac{\partial B}{\partial t} - k_\delta B + \frac{1}{2} \sigma_\delta^2 B^2 = 0,$$

$$\frac{\partial C}{\partial t} - C k_r + \frac{1}{2} \sigma_r^2 C^2 = 0,$$

$$B k_\delta \theta_\delta + C k_r \theta_r + \frac{\partial A}{\partial t} = 0,$$

with boundary condition of $A(u; u, T) = \alpha(u, T)$, $B(u; u, T) = \beta(u, T)$ and $C(u; u, T) = \gamma(u, T)$. Hence,

$$B(t; u, t) = \frac{2k_\delta}{\sigma_\delta^2 (1 + q_\delta e^{-k_\delta t})},$$

$$C(t; u, t) = \frac{2k_r}{\sigma_r^2 (1 + q_r e^{-k_r t})},$$

$$A(t; u, T)$$

$$\begin{aligned} &= \alpha(u, T) + \frac{2k_\delta \theta_\delta}{\sigma_\delta^2} \left[k_\delta (u - t) + \ln \left(\frac{1 + q_\delta e^{-k_\delta u}}{1 + q_\delta e^{-k_\delta t}} \right) \right] \\ &\quad + \frac{2k_r \theta_r}{\sigma_r^2} \left[k_r (u - t) + \ln \left(\frac{1 + q_r e^{-k_r u}}{1 + q_r e^{-k_r t}} \right) \right] \end{aligned}$$

with

$$q_\delta = \left[\frac{2k_\delta}{\beta(u, T) \sigma_\delta^2} - 1 \right] e^{k_\delta u},$$

$$q_r = \left[\frac{2k_r}{\gamma(u, T)\sigma_r^2} - 1 \right] e^{k_r u}.$$

Therefore,

$$F(S_t, \delta_t, t; T) = S_t e^{\alpha(t, T) + \beta(t, T)\delta_t + \gamma(t, T)r_t} + w \int_t^T e^{A(u; t, T) + B(u; t, T)\delta_t + C(u; t, T)r_t} du.$$

Appendix F. The alternative model with the convenience yield

following the Vasicek process

The alternative model can be expressed as follows. In the risk-neutral measure,

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_{1,t}^{\mathbb{Q}},$$

$$d\delta_t = \kappa_\delta(\theta_\delta - \delta_t)dt + \sigma_\delta\sqrt{r_t}dW_{2,t}^{\mathbb{Q}},$$

$$dS_t = [(r_t - \delta_t)S_t + w]dt + S_t\sigma_{Sr}\sqrt{r_t}dW_{1,t}^{\mathbb{Q}} + S_t\sigma_{S\delta}dW_{2,t}^{\mathbb{Q}} + S_t\sqrt{v_0 + v_{Sr}r_t}dW_{3,t}^{\mathbb{Q}},$$

$$dW_1^{\mathbb{Q}}dW_2^{\mathbb{Q}} = dW_1^{\mathbb{Q}}dW_3^{\mathbb{Q}} = dW_2^{\mathbb{Q}}dW_3^{\mathbb{Q}} = 0.$$

Note that in this model since δ can be negative, $v_{S\delta}$ must be zero in order to make the spot volatility admissible.

We assume the risk premium follows:

$$d \begin{bmatrix} W_{1,t}^{\mathbb{Q}} \\ W_{2,t}^{\mathbb{Q}} \\ W_{3,t}^{\mathbb{Q}} \end{bmatrix} = \begin{bmatrix} \eta_r\sqrt{r_t} \\ \eta_\delta \\ \eta_S\sqrt{v_0 + v_{Sr}r_t} \end{bmatrix} dt + d \begin{bmatrix} W_{1,t}^{\mathbb{P}} \\ W_{2,t}^{\mathbb{P}} \\ W_{3,t}^{\mathbb{P}} \end{bmatrix}.$$

Thus, in the physical measure, the process follows

$$dr_t = [k_r(\theta_r - r_t) + \sigma_r\eta_r r_t]dt + \sigma_r\sqrt{r_t}dW_{1,t}^{\mathbb{P}}$$

$$d\delta_t = [k_\delta(\theta_\delta - \delta_t) + \sigma_\delta\eta_\delta]dt + \sigma_\delta dW_{2,t}^{\mathbb{P}}$$

$$dS_t = [(r_t - \delta_t + \sigma_{Sr}\eta_r r_t + \sigma_{S\delta}\eta_\delta + \eta_S(v_0 + v_{Sr}r_t))S_t + w]dt + S_t\sigma_{Sr}\sqrt{r_t}dW_{1,t}^{\mathbb{P}} \\ + S_t\sigma_{S\delta}dW_{2,t}^{\mathbb{P}} + S_t\sqrt{v_0 + v_{Sr}r_t}dW_{3,t}^{\mathbb{P}}$$

$$dW_1^{\mathbb{P}}dW_2^{\mathbb{P}} = dW_1^{\mathbb{P}}dW_3^{\mathbb{P}} = dW_2^{\mathbb{P}}dW_3^{\mathbb{P}} = 0$$

References

- Alizadeh, A., Nomikos, N., Pouliasis, P., 2008. A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance* 32, 970–1983.
- Arnold, L., 1974. *Stochastic Differential Equations: Theory and Applications*. John Wiley & Sons, New York.
- Bessembinder, H., Chan, K., 1992. Time varying risk premia and forecastable returns in futures markets. *Journal of Financial Economics* 32, 169–193.
- Bessembinder, H., Coughenour, J., Seguin, P., Smoller, M., 1995. Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *Journal of Finance* 50, 361–375.
- Birkhoff, G., Gian-Carlo, R., 1989. *Ordinary Differential Equation*. John Wiley & Sons, New York.
- Brennan, M., 1958. The supply of storage. *American Economic Review* 48, 50–72.
- Brown, S., Dybvig, P., 1986. The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates. *Journal of Finance* 41, 617–630.
- Casassus, J., Collin-Dufresne, P., 2005. Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance* 60, 2283–2331.
- Chen, R., Scott, L., 1993. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income* 3, 14–31.
- Chen, R., Scott, L., 2003. Multi-factor Cox–Ingersoll–Ross models of the term structure: Estimates and tests from a Kalman filter model. *Journal of Real Estate Finance and Economics* 27, 143–172.
- Cox, J., Ingersoll, J., Ross, S., 1981. The relation between forward prices and futures prices. *Journal of Financial Economics* 9, 321–346.

- Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.
- Deaton, A., Laroque, G., 1992. On the behaviour of commodity prices. *Review of Economic Studies* 59, 1–23.
- Deaton, A., Laroque, G., 1996. Competitive storage and commodity price dynamics. *Journal of Political Economy* 104, 896–923.
- Dempster, M., Medova, E., Tang, K., 2008. Long term spread option valuation and hedging. *Journal of Banking and Finance* 32, 2530–2540.
- Duffee, G., 1999. Estimating the price of default risk. *Review of Financial Studies* 12, 197–226.
- Duffee, G., 2002. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57, 405–443.
- Duffee, G., Stanton, R., 2004. *Estimation of dynamic term structure models*. Working Paper, Haas School of Business, Univ. of California, Berkeley.
- Duffie, D., 2006. *Dynamic Asset Pricing*. Princeton University Press, Princeton.
- Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. *Mathematical Finance* 6, 379–406.
- Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusion. *Econometrica* 68, 1343–1376.
- Fama, E., French, K., 1988. Business cycles and the behavior of metals prices. *Journal of Finance* 43, 1075–1093.
- Gallant, R., Tauchen, G., 1996. Which moments to match? *Econometric Theory* 12, 657–681.
- Geman, H., Ohana, S., 2008. Time-consistency in managing a commodity portfolio: A dynamic risk measure approach. *Journal of Banking and Finance* 32, 1991–2005.

- Geman, H., Nguyen, V., 2005. Soybean inventory and forward curve dynamics. *Management Science* 51, 1076–1091.
- Gibson, R., Schwartz, E., 1990. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance* 45, 959–976.
- Gorton, G., Hayashi, F., Rouwenhorst, K., 2007. *The fundamentals of commodity futures returns*. Yale ICF Working Paper, No. 07–08.
- Hilliard, J., Reis, J., 1998. Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *Journal of Financial and Quantitative Analysis* 33, 61–86.
- Kaldor, N., 1939. Speculation and economic stability. *Review of Economic Studies* 7, 1–27.
- Kim, T., Omberg, E., 1996. Dynamics of nonmyopic portfolio behaviour. *Review of Financial Studies* 9, 141–161.
- Lewis, M., 2007. The universe of commodity indices. *Commodity Now* 11, 40–46.
- Lien, D., Yang, L., 2008. Asymmetric effect of basis on dynamic futures hedging: Empirical evidence from commodity markets. *Journal of Banking and Finance* 32, 187–198.
- Litzenberger, R., Rabinowitz, N., 1995. Backwardation in oil futures markets: Theory and empirical evidence. *Journal of Finance* 50, 1517–1545.
- Miltersen, K., Schwartz, E., 1998. Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis* 33, 33–59.
- Oksendal, B., 2005. *Stochastic Differential Equations*. Springer Press, New York.
- Pindyck, R., 2001. The dynamics of commodity spot and futures markets: A primer. *The Energy Journal* 22, 1–29.

- Routledge, B., Seppi, D., Spatt, C., 2000. Equilibrium forward curves for commodities. *Journal of Finance* 55, 1297–1338.
- Schwartz, E., 1997. The stochastic behaviour of commodity prices: Implications for valuation and hedging. *Journal of Finance* 52, 923–973.
- Vasicek, O., 1977. An equilibrium characterisation of the term structure. *Journal of Financial Economics* 5, 177–188.
- Working, H., 1949. The theory of the price of storage. *American Economic Review* 39, 1254–1262.

Table 1. Parameters of interest rates and NCYs used in different applications. We use the method in Gibson and Schwartz (1990) to obtain the implied convenience yield from copper futures.

Application	Interest rate	Copper NCY	Copper implied NCY
Source:	Federal reserve	Schwartz (1997) Table 7	This paper
Sample:	3 Month T-bill 1985-2006 (weekly)	1988–1995 (weekly)	2000–2006 (daily)
Mean (%):	4.9	2.65	−0.7
Volatility (%):	0.8	25	32.3
Boundary	0	Roughly −2%	Roughly −2%

Table 2. Probability of negative convenience yield and the volatility of the convenience yield from Vasicek process. The probability of negative convenience yield is obtained by simulation, the starting (net) convenience yield and interest rate are assumed to be their long-run mean. As a comparison, we calculate the probability of negative interest rate resulting from Vasicek model in column 2. In Model 1, we simulate using the parameters directly obtained from Schwartz (1997), where weekly data are used from 1990 to 1995 and from 1988 to 1995 for oil and copper, respectively. In model 2, we calibrate the Schwartz (1997) model using data in this paper (see ** Section 5.1), the estimated parameters are then used for calculation. In model 3 we calibrate the alternative model in ** Appendix F using our data.

T	Interest rate	Model 1		Model 2		Model 3	
		Oil	Copper	Oil	Copper	Oil	Copper
<i>Panel A: Probability of negativity from Vasicek process</i>							
3 months	0.0000	0.713	0.731	0.151	0.133	0.103	0.127
6 months	0.0000	0.787	0.806	0.278	0.286	0.120	0.156
12 months	0.0001	0.874	0.881	0.358	0.400	0.159	0.167
<i>Panel B: Volatility generated from Vasicek process</i>							
Volatility	0.0008	0.372	0.249	0.228	0.101	0.145	0.085

Table 3. Mean and standard deviation of oil and copper returns. Daily data from January 03 2000 to September 01 2006 (1665 observations for each commodity) are used in the calibration. The futures prices of WTI crude oil (*CL*) and high-grade copper (*HG*) are obtained from the NYMEX. F_n denotes the n th futures contract closest to maturity.

	F_1	F_5	F_9	F_{13}	F_{17}
<i>Oil</i>					
Mean (daily return %)	0.086	0.086	0.088	0.090	0.092
Std (daily return %)	2.280	1.794	1.596	1.491	1.430
<i>Copper</i>					
Mean (daily return %)	0.098	0.095	0.092	0.090	0.087
Std (daily return %)	1.5877	1.5251	1.4615	1.4384	1.4431

Table 4. LME average warehouse monthly storage cost (USD per short ton). The storage cost is for the period of April 2007 to March 2008.

Aluminium alloy	Lead	NASAAC	Nickel	Primary Aluminium	Tin	Zinc
10.81	7.97	11.20	10.67	9.44	9.82	8.51

Table 5. Parameter estimation for oil and copper contracts. Numbers in brackets are standard deviations. Daily data from January 03 2000 to September 01 2006 are used in the calibration respectively, for oil and copper.

Variables	Oil	Copper
κ_r	0.1524(0.0130)	0.1057(0.0105)
θ_r	0.0997(0.0061)	0.1268(0.0097)
σ_r	0.0407(0.0007)	0.0404(0.0007)
κ_δ	0.4458(0.0087)	0.3105(0.0094)
θ_δ	0.2230(0.0025)	0.0943(0.0027)
σ_δ	0.2524(0.0049)	0.2640(0.0056)
σ_{Sr}	-0.1021(0.0417)	-0.0628(0.030)
$\sigma_{S\delta}$	0.2886(0.0144)	0.3575(0.0219)
ν_{Sd}	0.0495(0.0218)	0.2567(0.0226)
ν_{Sr}	0.3316(0.1762)	0.0187(0.411)
ν_0	0.0556(0.0066)	0.0257(0.0017)
η_r	-2.1189(1.2152)	-0.2819(0.1806)
η_δ	0.4030(0.0542)	0.6914(0.1121)
η_S	1.9756(0.3102)	4.7990(1.7385)
ξ_{F1}	0.0337(0.0064)	0.0161(0.0031)
ξ_{F2}	0.0086(0.0016)	0.0006(0.0002)
ξ_{F3}	0.0000(0.0000)	0.0027(0.0005)
ξ_{F4}	0.0010(0.0001)	0.0000(0.0000)
ξ_{F5}	0.0042(0.0008)	0.0072(0.0014)
ξ_{r1}	0.0000(0.0000)	0.0000(0.0000)
ξ_{r2}	0.0014(0.0003)	0.0014(0.0003)
Quasi-loglikelihood	47,570	49,581

Table 6. The put calendar spread option values of our model and the Schwartz model in four different days. The spread option values are obtained through Monte Carlo simulation (1000 paths with anti-variate techniques for each model).

First (January 03, 2000) $F_{13} = 20.00, F_{17} = 19.25, \zeta(t_0) = 0.75$							
Strike	$\zeta(t_0) - 1.5$	$\zeta(t_0) - 1$	$\zeta(t_0) - 0.5$	$\zeta(t_0)$	$\zeta(t_0) + 0.5$	$\zeta(t_0) + 1$	$\zeta(t_0) + 1.5$
Schwartz	0.0049	0.0140	0.0479	0.1973	0.5197	0.2496	1.4088
Our model	0.0011	0.0063	0.0529	0.2201	0.5121	0.2687	1.3132
500 th (January 07, 2002) $F_{13} = 21.6, F_{17} = 21.56, \zeta(t_0) = 0.04$							
Strike	$\zeta(t_0) - 1.5$	$\zeta(t_0) - 1$	$\zeta(t_0) - 0.5$	$\zeta(t_0)$	$\zeta(t_0) + 0.5$	$\zeta(t_0) + 1$	$\zeta(t_0) + 1.5$
Schwartz	0.0216	0.0450	0.1002	0.2410	0.5388	0.2860	1.4557
Our model	0.0005	0.0025	0.0240	0.1527	0.4551	0.1999	1.3342
1000 th (January 08, 2004) $F_{13} = 29.26, F_{17} = 28.44, \zeta(t_0) = 0.82$							
Strike	$\zeta(t_0) - 1.5$	$\zeta(t_0) - 1$	$\zeta(t_0) - 0.5$	$\zeta(t_0)$	$\zeta(t_0) + 0.5$	$\zeta(t_0) + 1$	$\zeta(t_0) + 1.5$
Schwartz	0.0270	0.0541	0.1258	0.3015	0.6062	1.0175	1.4811
Our model	0.0006	0.0040	0.0463	0.2405	0.6021	1.0727	1.6089
Last (September 01, 2006) $F_{13} = 74.41, F_{17} = 74.35, \zeta(t_0) = 0.06$							
Strike	$\zeta(t_0) - 3$	$\zeta(t_0) - 2$	$\zeta(t_0) - 1$	$\zeta(t_0)$	$\zeta(t_0) + 1$	$\zeta(t_0) + 2$	$\zeta(t_0) + 3$
Schwartz	0.1869	0.2961	0.4870	0.8079	1.3105	2.0078	2.8466
Our model	0	0.0005	0.0090	0.1711	0.7082	1.4930	2.3832

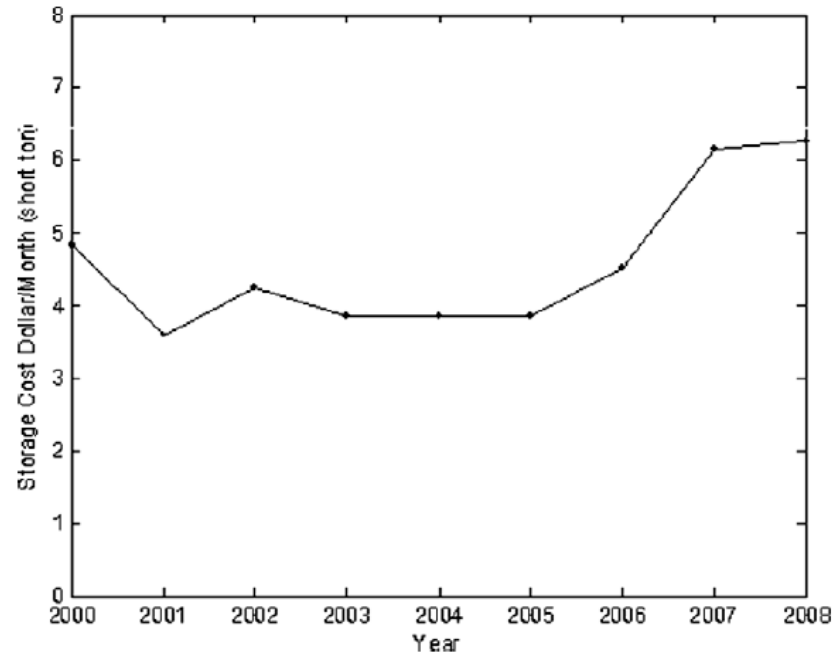


Figure 1. Monthly storage cost for copper. The data are obtained from NYMEX from 2000 to 2008. The storage cost is re-announced annually.